四种重要医学媒介蚊虫离子受体基因 *IR8a* 和 *IR25a* 的特征及分类地位

王婷婷，郝友进，何正波，梅 婷，陈 斌*

（重庆师范大学昆虫与分子生物学研究所，媒介昆虫重庆市重点实验室，重庆 401331）

摘要：【目的】鉴定中华按蚊 *Anopheles sinensis*、冈比亚按蚊 *Anopheles gambiae*、埃及伊蚊 *Aedes aegypti* 和致倦库蚊 *Culex quinquefasciatus* 4 种重要医学媒介蚊虫的离子受体基因 *IR8a* 与 *IR25a* 以及代表性的 iGluRs 基因和 IRs 基因，研究这些基因的特征及其系统发育关系，确定 *IR8a* 与 *IR25a* 基因在离子谷氨酸受体基因中的分类地位。【方法】以黑腹果蝇 *Drosophila melanogaster* 的 iGluRs 和 IRs 基因序列作为询问序列，用 BLASTP 和 BLASTN 方法分别搜索中华按蚊、冈比亚按蚊、埃及伊蚊和致倦库蚊的基因组数据库，以鉴定这 4 个种内的 iGluRs 和 IRs 基因序列；用生物信息学方法比较分析鉴定获得的 *IR8a* 与 *IR25a* 及代表性的 iGluRs 和 IRs 基因氨基酸序列的特征；用最大似然法、贝叶斯法和最大简约法构建并讨论这些基因氨基酸序列的系统发育关系；用 PAML 软件计算这些基因的 $d_N/d_S(\omega)$ 值并分析其选择压力。【结果】通过对中华按蚊、冈比亚按蚊、埃及伊蚊和致倦库蚊中离子谷氨酸受体基因进行生物信息学鉴定，获得了所有 *IR8a* 和 *IR25a* 基因，24 个代表性的 iGluRs 基因和 30 个代表性的 IRs 基因的序列。特征比较和系统发育关系分析结果表明，在 4 种蚊虫和黑腹果蝇中 *IR8a* 与 *IR25a* 基因的氨基酸序列长度与 iGluRs 的更接近，其平均值远大于其他 IRs 的；IR25a 和 IR8a 有与传统 iGluRs 相同的氨基酸末端区域（amino-terminal domain, ATD），虽然 IR8a 的序列不太保守，而 IRs 则没有 ATD；IR8a 与 IR25a 序列有与传统 iGluRs 一样保守的配体结合区 S1 和 S2，而 IRs 没有。所有这些基因的 ω 值远小于 1，意味着这些基因在进化过程中都经历了纯化选择，而 *IR8a* 与 *IR25a* 的 ω 值更接近 iGluRs。此外，IR8a 与 IR25a 还有与 non-NMDA 相似的特异性位点，而这些位点在 IRs 不存在。再者，IR8a 和 IR25a 与 non-NMDA 聚在一枝，形成一个姐妹群，群内序列间的遗传距离小于本研究中任何其他序列间的距离。根据这些研究结果，我们将 IR8a 与 IR25a 分类进入传统 iGluRs 中的 non-NMDA 类型，并命名为 Putative 受体。【讨论】本研究构建了蚊虫 iGluRs 和 IRs 基因的信息框架，确定了 *IR8a* 与 *IR25a* 基因的分类地位，对进一步的功能研究具有重要意义。

关键词：蚊虫；iGluRs；IRs；*IR8a*；*IR25a*；全基因组鉴定；序列特征；系统发育

中图分类号：Q966　　**文献标识码**：A　　**文章编号**：0454-6296(2017)04-0379-10

Characteristics and classification position of the ionotropic receptor genes *IR8a* and *IR25a* in four vector mosquito species of medical importance

WANG Ting-Ting, HAO You-Jin, HE Zheng-Bo, MEI Ting, CHEN Bin* (Chongqing Key Laboratory of Vector Insects, Institute of Entomology and Molecular Biology, Chongqing Normal University, Chongqing 401331, China)

Abstract: 【Aim】 To identify the ionotropic receptor genes *IR8a* and *IR25a*, and the representative iGluRs and IRs genes in four vector mosquito species of medical importance, *Anopheles sinensis*, *Anopheles gambiae*, *Aedes aegypti* and *Culex quinquefasciatus*, and to investigate the characteristics and

基金项目："两江学者"计划专项经费；国家自然科学基金项目(31672363, 31372265)；国际原子能机构 CRP 项目(18268)；国家科技基础性工作专项重点项目(2015FY210300)

作者简介：王婷婷，女，1993 年 1 月生，山西长治人，硕士研究生，研究方向为昆虫分子生物学，E-mail: 1210539940@qq.com

* 通讯作者 Corresponding author, E-mail: bin.chen@cqnu.edu.cn

收稿日期 Received: 2017-01-18；接受日期 Accepted: 2017-03-07

phylogenetic relationships of these genes, so as to determine the classification position of the *IR8a* and *IR25a* genes in the group of ionotropic glutamate receptor genes. 【Methods】By using BLASTP and BLASTN searching against the genomes of these four mosquito species with the iGluRs and IRs gene sequences of *Drosophila melanogaster* as query sequences, the sequences of iGluRs and IRs genes of these four mosquito species were identified. The characteristics of the amino acid sequences of *IR8a* and *IR25a* and the representative iGluRs and IRs genes identified were predicted and comparatively analyzed by bioinformatics methods. The phylogenetic relationships of these genes in the five species were deduced and discussed by using maximum likelihood, Bayesian and maximum parsimony methods based on amino acid sequences, and the selection pressures of these genes were analyzed with the $d_N/d_S(\omega)$ values calculated by using PAML software. 【Results】The sequences of *IR8a* and *IR25a*, 24 representative traditional iGluRs genes and 30 representative IRs genes in mosquitoes *An. sinensis*, *An. gambiae*, *An. aegypti* and *Cx. quinquefasciatus* were identified by bioinformatics methods. The comparative analyses of characteristics and phylogenetic relationships of these genes showed that the amino acid sequence lengths of *IR8a* and *IR25a* in the four mosquito species and *D. melanogaster* are closer to those of iGluRs, and their mean values are significantly longer than those of other IRs. IR25a has the amino-terminal domain (ATD) that is the characteristic of traditional iGluRs, and IR8a also has a region similar to ATD, although it is not conservative, while IRs have no ATD. IR8a and IR25 also have the conserved S1 and S2 ligand-binding domain (LBD), the characteristics of iGluRs, but IRs do not have it. The $d_N/d_S(\omega)$ values of all these genes were far less than 1, suggesting the purifying selection in evolution, while the $d_N/d_S(\omega)$ values of *IR8a* and *IR25a* were closer to those of iGluRs. In addition, like non-NMDA, IR8a and IR25a have specific sites which do not exit in IRs. IR8a and IR25a cluster together with non-NMDA to form a sister group, with the genetic distance between the sequences within the group less than that of any other pairs of sequences in this study. Based on these results, IR8a and IR25a were classified into the non-NMDA receptors in iGluRs, and named as Putative receptors. 【Conclusion】This study developed the information frame of iGluRs and IRs genes of mosquito species and determined the classification position of *IR8a* and *IR25a* genes, which is of significance for further functional study of these genes.

Key words: Culicidae; iGluRs; IRs; *IR8a*; *IR25a*; genome-wide identification; sequence characteristics; phylogeny

离子谷氨酸受体(ionotropic glutamate receptors, iGluRs)是真核生物和原核生物中一类保守的配体结合离子通道,在脊椎动物神经系统突触间的信号传递中作为兴奋的神经递质谷氨酸的受体,通过调节阳离子流跨质膜来应答细胞外谷氨酸和相关配体的结合,细胞可以利用其来调节内部生理以应对外部环境信号(Mayer and Armstrong, 2004)。iGluRs在药理学和分子特征上被分为3个亚类,根据它们的主要激动剂将其分别命名为:α-氨基-3羟基-5甲基-4异恶唑受体(AMPA receptors)、红藻氨酸受体(Kainate receptors)和N-甲基-D-天冬氨酸受体(NMDA receptors)。AMPA受体主要在大部分中枢神经系统突触内调节快速兴奋性突触传导,而Kainate受体在这个过程中只有微弱的调节作用,它们通常被合称为non-NMDA受体。NMDA受体则在突触的可塑性(synaptic plasticity)中发挥功能,需要谷氨酸和甘氨酸两种激活剂来激活(Heidelberg, 2004; Ashby *et al*., 2008)。

离子受体(ionotropic receptors, IRs)家族是iGluRs中新发现的一类基因(Benton *et al*., 2009),与嗅觉相关(Vosshall and Stocker, 2007; Masse *et al*., 2009),在既不表达嗅觉受体(olfactory receptor, OR)基因也不表达味觉受体(gustatory receptor, GR)基因的触角腔锥形神经元中表达(Vosshall and Stocker, 2007; Montell, 2009),因此被认为是一类新的化学感应受体基因,在检测外界环境中的挥发性化学物质以及细胞间化学信号的过程中有广泛的功能(Benton, 2007)。传统的iGluRs和IRs已在许多动物物种中被鉴定出来(Tikhonov and Magazanik, 2009)。在黑腹果蝇*Drosophila melanogaster*、冈比亚按蚊*Anopheles gambiae*、埃及伊蚊*Aedes aegypti*和致倦库蚊*Culex quinquefasciatus*基因组中,分别有12, 9, 8和6个iGluRs基因及66, 46, 94和69个IRs基因(Liu *et al*., 2010; Croset *et al*., 2010)。在所有

已鉴定 IRs 基因的昆虫中，都有 *IR8a* 与 *IR25a* 基因存在，它们在触角的大部分位置单独或与别的 IR 基因一起表达（Benton *et al.*, 2009），被归类为 IRs 基因。但是由于 *IR8a* 与 *IR25a* 基因在序列特征、保守性和系统发育关系等方面的特异性，它们在分类家族的归属上一直是一个未解决的问题（Croset *et al.*, 2010）。

首先在氨基酸序列上，传统 iGluRs（NMDA, non-NMDA）和 IRs 都包含一个胞外 N-末端，一个被离子通道孔 P 隔开的由 S1 和 S2 构成的配体结合区（ligand binding domain, LBD）（Mayer and Armstrong, 2004; Jin *et al.*, 2009），3 个跨膜区（M1 – M3）和一个胞内的 C-末端（Mayer *et al.*, 2006）。其中离子通道孔 P 和跨膜区 M1 – M3 组成离子通道结构域 Lig_Chan（PF00060）（Kuner *et al.*, 2003）。此外，传统的 iGluRs 在 N 末端还有一个氨基酸末端区域（amino-terminal domain, ATD）（PF01094），负责通道的组装以及结合辅助因子。黑腹果蝇的 IR8a 和 IR25a 也有与 ATD 相似的序列，只是更分散（Croset *et al.*, 2010），然而其他的 IRs（不包括 IR8a 和 IR25a，下同）则没有这个序列。其次，传统 iGluRs（不包括 IRs）的配体结合位点保守性强，它们在 S1 区都有与谷氨酸盐的 α 羧基直接结合的精氨酸（R），S2 区都有与谷氨酸盐的 γ 羧基直接结合的苏氨酸（T）和与谷氨酸盐的 α 氨基酸直接结合的天冬氨酸（D）或谷氨酸（E）（Naur *et al.*, 2007）。黑腹果蝇、豌豆蚜 *Acyrthosiphon pisum*、赤拟谷盗 *Tribolium castaneum*、灰翅夜蛾 *Spodoptera littoralis* 和家蚕 *Bombyx mori* 中的 IR8a 同时有 3 个配体结合位点，IR25a 也有其中两个（T, D/E）。其他 IRs 的配体结合区则非常分散，大部分都没有或者只有其中一个位点（Benton *et al.*, 2009; 郭梅, 2014），可能结合其他类型的配体行使不同的功能。再者，在黑腹果蝇中基于氨基酸序列构建的 iGluRs, IRs, IR8a 和 IR25a 的系统发育上，IR8a 与 IR25a 不聚集在 IR 家族支上，而在 iGluRs 支上（Bengtsson *et al.*, 2012; Liu *et al.*, 2012），进化速率也与 iGluRs 更接近，远慢于其他的 IRs 基因。这种情况在螟虫和棉铃虫 *Helicoverpa armigera* 等物种的系统发育关系中也有发现（Liu *et al.*, 2012; Glaser *et al.*, 2013）。

本研究基于重庆师范大学中华按蚊基因组和转录组测序数据（Chen *et al.*, 2014）及从 Vectorbase 下载的冈比亚按蚊、埃及伊蚊和致倦库蚊的基因组数据，以黑腹果蝇 iGluRs 和 IRs 基因序列作为询问序列，采用 BLASTN 和 BLASTP 方法，从信息学上鉴定了这 4 种重要医学媒介蚊虫的 *IR8a* 与 *IR25a* 基因以及代表性的 iGluRs 和 IRs 基因，并进行一系列特征分析。通过将 IR8a 和 IR25a 与 iGluRs 和 IRs 从氨基酸序列长度、结构域、保守性、替换率、特异性位点和系统发育关系 6 个方面进行比较分析，确认了 IR8a 与 IR25a 的特殊性，并将它们归为 iGluRs 中的 non-NMDA 受体类型，命名为 Putative 受体。此研究丰富了蚊虫的 iGluRs 基因数据，为进一步研究 *IR8a* 与 *IR25a* 基因提供了信息基础。

1　材料与方法

1.1　数据来源

中华按蚊基因组数据包括基因组序列和注释以及预测的基因集序列，由重庆师范大学昆虫和分子生物学研究所在华大基因完成，组装覆盖度达到了转录组测序产生的蛋白编码 ESTs 的 98.35%（Chen *et al.*, 2014）。冈比亚按蚊、埃及伊蚊和致倦库蚊的基因组数据库和黑腹果蝇的 iGluRs、*IR8a* 和 *IR25a* 以及 IRs 基因序列下载自 Vectorbase（https://www.vectorbase.org/）。

1.2　基因的生物信息学鉴定

首先，以黑腹果蝇 iGluRs 和 IRs 基因作为查询序列，对上述 4 种蚊虫的基因组数据库进行 BLASTP 和 BLASTN 同源性搜索，E-value $< 1 \times 10^{-5}$ 作为阈值，将得到的候选基因在 Softberry 中预测其转录组和蛋白序列后，于 NCBI 中再进行 BLASP 搜索，查看其结构域以及与查询序列的相似度，鉴定得到正确的基因。

1.3　基因特征分析

使用 BLAST 工具（http://www.ncbi.nlm.nih.gov/BLAST/）进行序列相似性搜索；使用 NCBI Blast 和 Interproscan（http://www.ebi.ac.uk/interpro/）对氨基酸序列进行结构域预测和亚细胞定位。使用 Clustal W 软件（Thompson *et al.*, 2002）对 4 种蚊虫和黑腹果蝇的代表性 iGluRs, IR8a, IR25a 和 IRs 的氨基酸序列进行多重比对。Gap Opening Penalty 参数设为 10，Gap Extension Penalty 设为 0.2；Gap Separation Distance 设为 4；Delay Divergent Cutoff（%）设为 30%。蛋白质的三级结构，要首先通过 PSI-BLAST 在 PDB（Protein Data Bank）中搜索得到氨基酸序列一致性最高（大于 80%），且具有相应结构域的基因，用其作为同源建模的模板，再采用 SWISS-MODEL（http://swissmodel.expasy.org/）进

行同源建模和 3D 结构预测(唐尧等, 2014)。

1.4 系统发育分析

将中华按蚊、冈比亚按蚊、致倦库蚊、埃及伊蚊和黑腹果蝇 5 种昆虫的 IR8a 和 IR25a (共 10 个)及代表性的 iGluRs(24 个)和代表性的 IRs (30 个)氨基酸序列,用 Clustal W 对其进行多重比对,分别用最大似然法(maximum likelihood, ML)、贝叶斯法(Bayesian inference, BI)和最大简约法(maximum parsimony, MP)构建系统发育树。其中,用 Mega 5.0 软件构建 ML 和 MP 系统发育树,用 Find Best Protein Model 进行模型预测, ML 树用 WAG 模型。使用 1 000 次重复计算系统发育树上的 bootstrap 值(自展分析值)。使用 PAUP 软件构建贝叶斯系统发育树,用 Modeltest3.7 软件分析模型, BI 树也用 WAG 模型。大于 50% 的自展值标于系统发育树相应分支上,讨论其系统发育关系。

1.5 基因替换率分析

使用 PAML(Yang, 1997)软件中的最大似然法计算基因的 $d_N/d_S(\omega)$ 值。以中华按蚊的序列作为对照,用 MEGA 软件中的 MUSCLE codon 分别将其与冈比亚按蚊,致倦库蚊和埃及伊蚊直系同源基因的编码区(CDs)进行比对,手动删除其中的终止子,存为 FASTA 格式。将 FASTA 格式转换成 PAML 格式后,输入 PAML 软件,于 PAML 的 yn00 中,用 M0 模型计算非同义突变频率(d_N)、同义突变频率(d_S)和 $d_N/d_S(\omega)$的值, $\omega \ll 1$ 为纯化选择, $\omega \gg 1$ 为正向选择, $\omega = 1$ 为中性选择(Swanson and Vacquier, 2002; Swanson *et al.*, 2003)。本研究计算了 3 种蚊虫中的 *IR8a* 和 *IR25a* 基因及代表性 iGluRs 和 IRs 基因的 ω 值,比对分析所受到的选择压力和保守性。

2 结果

2.1 基因的生物信息学鉴定

本研究通过对中华按蚊、冈比亚按蚊、埃及伊蚊和致倦库蚊中的离子谷氨酸受体基因进行生物信息学鉴定,获得了所有 *IR8a* 和 *IR25a* 基因,以及 24 个代表性的 iGluRs 基因和 30 个代表性的 IRs 基因的序列信息。其中致倦库蚊 *CqIR21a*, *CqIR25a* 和 *CqIR68a* 是非全长序列,因此没有包括在后续 3 类基因氨基酸序列长度和 $d_N/d_S(\omega)$值的比较计算中。

2.2 *IR8a* 与 *IR25a* 基因的结构

基于 4 种蚊虫和黑腹果蝇 *IR8a* 和 *IR25a* 基因组序列和转录组序列,在 Gene Structure Display Server 在线网站中预测得到如图 1 所示的基因结构。*IR8a* 基因有 3 ~ 5 个外显子(图 1: A), *IR25a* 基因有 8 ~ 10 个外显子(图 1: B)。对于 *IR8a* 和 *IR25a* 两个基因的内含子,长度在 50 ~ 100 bp 的个数占 58.18%,长度在 100 ~ 200 bp 的个数占 21.82% (图 2);0, 1 和 2 相位内含子所占比例分别为 30.91%, 34.55% 和 34.55%。*AaIR8a* 基因的结构与其他基因差别较大,有一个大于 1 000 bp 的内含子(图 1: A)。

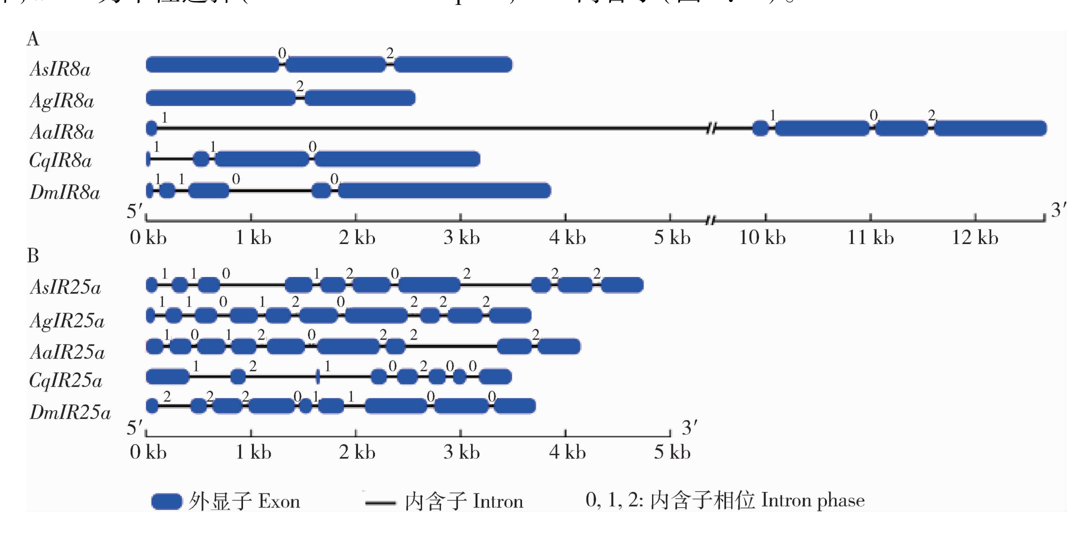


图 1 5 种昆虫 *IR8a* 与 *IR25a* 的基因结构

Fig. 1 Gene structure of *IR8a* and *IR25a* in five insect species

As: 中华按蚊 *Anopheles sinensis*; *Ag*: 冈比亚按蚊 *Anopheles gambiae*; *Aa*: 埃及伊蚊 *Aedes aegypti*; *Cq*: 致倦库蚊 *Culex quinquefasciatus*; *Dm*: 黑腹果蝇 *Drosophila melanogaster*. 外显子用蓝色的圆角矩形表示,内含子用黑色的线段表示,线上的数字表示各内含子的相位。Exons are marked with blue rounded rectangles, introns are marked with black lines, and the numbers above the black lines indicate intron phase.

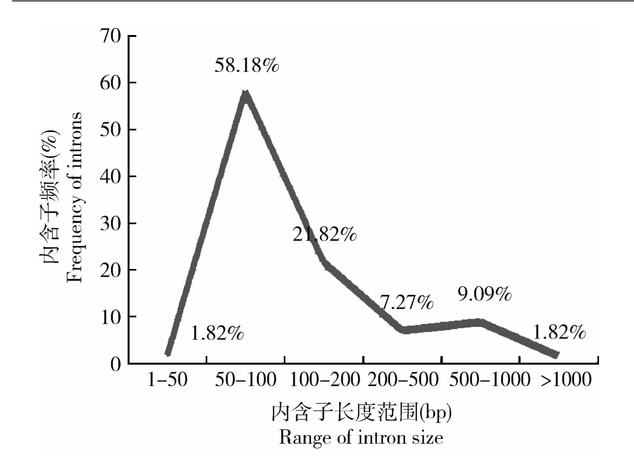


图 2　5 种昆虫 *IR8a* 与 *IR25a* 基因不同长度内含子的频率

Fig. 2　Frequencies of different lengths of introns of *IR8a* and *IR25a* in five insect species

2.3　iGluRs 和 IRs 氨基酸序列长度比较

4 种蚊虫和黑腹果蝇 iGluRs 的氨基酸序列长度在 632 ~ 1 055 之间(平均值 = 946 a. a.)；*IR8a* 与 *IR25a* 的氨基酸序列长度在 887 ~ 1 116 之间(平均值 = 940 a. a.)，而 IRs 的氨基酸序列较短，在 481 ~ 1 047 之间(平均值 = 667 a. a.)，表明 *IR8a* 和 *IR25a* 基因在氨基酸序列长度上与 iGluRs 的平均值更为接近(图 6：A)。

2.4　iGluRs 和 IRs 氨基酸序列结构域、配体结合位点和特异性位点比较

2.4.1　结构域比较：传统 iGluRs 和 IR25a(不包括非全长 CqIR25a)氨基酸序列都有氨基酸末端区域(ATD)、配体结合区(S1, S2)和离子通道区 Lig_Chan 3 个结构域；IR8a 有配体结合区和 Lig_Chan 结构域；而 IRs 则只有 Lig_Chan 结构域(图 3：B)。

2.4.2　配体结合位点比较：传统的 iGluRs 与 IR8a 氨基酸序列在配体结合区都有直接与谷氨酸配体结合的 3 个氨基酸位点：精氨酸(R)、苏氨酸(T)、天冬氨酸(D)或谷氨酸(E)。IR25a 有精氨酸(R)和天冬氨酸(D)两个氨基酸位点，IRs 无这 3 个位点(图 3：A)。这说明 IR8a 与 IR25a 的配体结合位点保守性强，不同于配体结合区极度分散的 IRs(Glaser *et al.*, 2013；Liu *et al.*, 2014)。

2.4.3　特异性位点比较：传统 iGluRs 的氨基酸序列有一些特异性位点(special sites)，在动物神经元突触内以及谷氨酸激活钾离子通道的过程中发挥作用。这些特异性位点有两个主要的特征：多肽结合位点(peptide binding site)(Brockie and Maricq, 2003)和二聚体界面(dimer interface)(Madden, 2002)，具有这些位点的氨基酸序列属于胞质结合蛋白 2 超家族的成员。Kainate 和 AMPA 受体的特异性位点分别是 PBP2_iGluRs_Kainate (cd13714)和 PBP2_iGluRs_AMPA (cd13715)，NMDA 受体的特异性位点是 PBP2_iGluRs_NMDA，根据组成亚基的不同又分为 NMDA_Nr1/Nr2/Nr3 (cd13719/cd06378/cd13720) 3 种。我们在 IR8a 与 IR25a 氨基酸序列中发现了 PBP2_iGluRs_putative (cd13717)特异性位点，而在 IRs 氨基酸序列中没有发现任何特异性位点。

通过氨基酸序列多重比对发现，上述所有这些特异性位点都出现在配体结合区 S1 和 S2 中，IR8a 和 IR25a 单独与传统 non-NMDA 受体相似的位点有 12 个(序列上方红色三角形标记的位点)，同时与传统 non-NMDA 和 NMDA 受体相似的位点有 11 个(序列上方蓝色圆形标记的位点)，但是没有单独与 NMDA 受体相似的位点(图 3：A)。另外我们在胞质结合蛋白 2 超家族的分层 CD 树(http://www.ncbi.nlm.nih.gov/Structure/cdd/cddsrv.cgi)中可以看到，具有 PBP2_iGluRs_Kainate, PBP2_iGluRs_AMPA 和 PBP2_iGluRs_putative 特异性位点的氨基酸序列都归类为 non-NMDA 受体(Brockie and Maricq, 2003)。因此从特异性位点可以看出，IR8a 和 IR25a 更接近传统 iGluRs 中的 non-NMDA 受体，建议将它们归为 non-NMDA 类型，并命名为 Putative 受体。

2.5　iGluRs 和 IRs 三级结构比较

以中华按蚊 AsGluRIIa, AsIR25a, AsIR8a 和 AsIR60a 为代表比较传统 iGluRs, IR25a, IR8a 和 IRs 氨基酸序列的三级结构(图 4)。AsGluRIIa 和 AsIR25a 氨基酸序列的三级结构从 N 末端到 C 末端依次为氨基酸末端区域(ATD)、配体结合区(LBD)和离子通道区(IC)。其中 ATD 以两边对称的形式组织，两个“近端的”ATD 以轴对称线形成一个小的分界面，两个“远端的 ATD”彼此并不接触；LBD 是由两个部分(S1,S2)构成的“捕蝇草”结构；而 IC 则主要由 α 螺旋组成，呈 4 倍折叠对称。AsIR8a 的三级结构除了有完整的 LBD 和 IC 外，其 N 末端也有 ATD 的结构，只是更分散；而 IR60a 则只有离子通道区完整的 IC 和分散的 LBD 区，没有 ATD 区域。结果表明，IR8a 与 IR25a 的三级结构与传统的 iGluRs 更接近。

2.6　iGluRs 和 IRs 的系统发育关系

将中华按蚊、冈比亚按蚊、埃及伊蚊、致倦库蚊和黑腹果蝇5种昆虫的IR8a和IR25a(共10个序

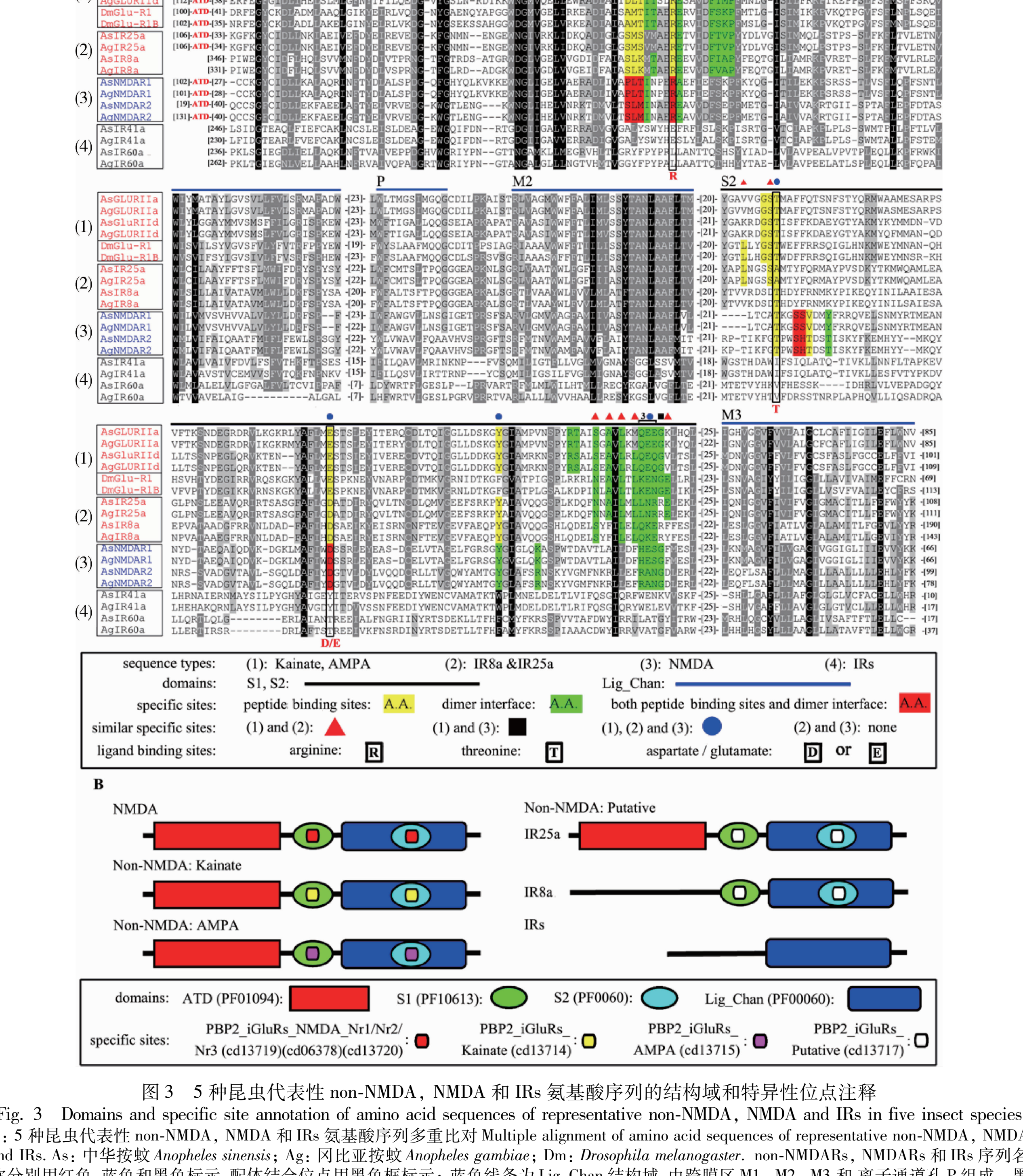


图 3 5 种昆虫代表性 non-NMDA，NMDA 和 IRs 氨基酸序列的结构域和特异性位点注释

Fig. 3 Domains and specific site annotation of amino acid sequences of representative non-NMDA, NMDA and IRs in five insect species

A：5 种昆虫代表性 non-NMDA，NMDA 和 IRs 氨基酸序列多重比对 Multiple alignment of amino acid sequences of representative non-NMDA, NMDA and IRs. As：中华按蚊 *Anopheles sinensis*；Ag：冈比亚按蚊 *Anopheles gambiae*；Dm：*Drosophila melanogaster*. non-NMDARs，NMDARs 和 IRs 序列名称分别用红色、蓝色和黑色标示，配体结合位点用黑色框标示；蓝色线条为 Lig_Chan 结构域，由跨膜区 M1，M2，M3 和 离子通道孔 P 组成。黑色、灰色和白色阴影的氨基酸序列其一致性分为别为 100%，70% ~99% 和 70% 以下。The names of amino acid sequences of non-NMDARs, NMDARs and IRs are in red, blue and black, respectively, and the black-lined boxes are ligand-gated sites. Lig_chan domain is marked with blue lines above, consisting of transmembrane region M1, M2 and M3 and ion channel pore P. The amino acid sequences with 100%, 70% –99% and below 70% identity are denoted with black, grey and white shades, respectively. B：5 种昆虫代表性 non-NMDA，NMDA 和 IRs 结构域和特异性位点 The features of domains and specific sites of representative non-NMDA, NMDA and IRs of five insect species. ATD：氨基酸末端区域 Amino-terminal domain；S1，S2：配体结合区 Ligand-gated sites；Lig_Chan：离子通道区 Ion channel region. NMDA 受体的特异性位点：PBP2_iGluRs_NMDA，包括 NMDA_Nr1/Nr2/Nr3（cd13719/cd06378/cd13720）3 种；non-NMDA 受体的特异性位点：Kainate 受体的特异性位点是 PBP2_iGluRs_Kainate（cd13714），AMPA 受体的特异性位点是 PBP2_iGluRs_AMPA（cd13715），Putative 受体的特异性位点是 PBP2_iGluRs_Putative（cd13717）。The specific sites for NMDA iGluRs: PBP2_iGluRs_NMDA with NMDA_Nr1/Nr2/Nr3 (cd13719/cd06378/cd13720) (three types); the specific sites for non-NMDA iGluRs: PBP2_iGluRs_Kainate (cd13714) for Kainate, PBP2_iGluRs_AMPA (cd13715) for AMPA, and PBP2_iGluRs_Putative (cd13717) for Putative receptors.

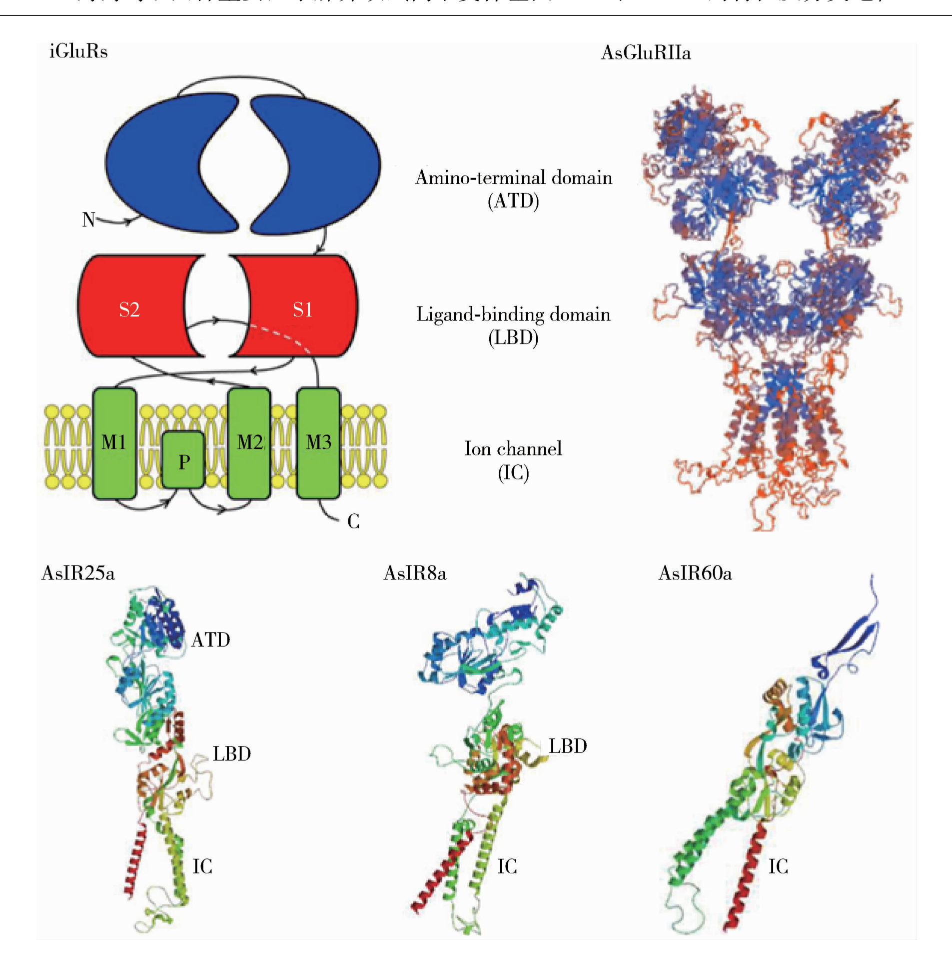


图 4　传统 iGluRs，IR25a，IR8a 和 IRs 代表性氨基酸序列的三维结构图

Fig. 4 Predicted 3D structure of representative amino acid sequences of traditional iGluRs，IR25a，IR8a and IRs

ATD：氨基酸末端区域 Amino-terminal domain；LBD：配体结合区 Ligand-binding domain；IC：离子通道区 Ion channel.

列），代表性的传统 iGluRs（24 个）及代表性 IRs（30 个）的氨基酸序列，分别用最大似然法（ML），贝叶斯（BI）和最大简约法（MP）构建系统发育树，结果表明这几种方法得到的结果基本一致。ML 系统发育树分析表明（图 5），各节点的自展值（bootstrap value）大多在 50% 以上，可信度高（Wiens and Servedio，1998），与 BI 和 MP 进化树一致，黑腹果蝇 non-NMDA 受体中出现了旁系同源受体，包括 AMPA 受体中的 DmGlu-R1B 和 DmGluR1，Kainate 受体中的 DmCG9935 和 DmCG11155，以及 DmGluRIId 和 DmCG5621。其他受体在不同物种中则都为直系同源受体。IR8a 和 IR25a 与传统 iGluRs 的 non-NMDA 受体聚在一支，它们与 NMDA 受体构成一个大的分支。3 种进化树的不同点在于，ML 树与 BI 树中，IRs 形成了一个单独的支系，而 MP 树中的 IRs 基因则形成了 3 个枝系。另外，从 ML 树中可以看到，IR8a 和 IR25a 与 non-NMDA 和 NMDA 受体间的遗传距离明显短于与 IRs 间的距离。这些从系统发育关系和遗传距离说明，IR8a 和 IR25a 与传统 iGluRs 中的 non-NMDA 受体最接近。

2.7　基因保守性和替换率的比较分析

本研究以中华按蚊传统 iGluRs 基因、*IR8a* 和 *IR25a* 基因，以及 IRs 基因作为对照，计算了冈比亚按蚊、埃及伊蚊和致倦库蚊中这 3 类基因的 d_N/d_S 比值 ω。如图 6（B）所示，传统 iGluRs 基因，*IR8a* 和 *IR25a* 基因，以及 IRs 基因这 3 类基因的 ω 值都远小于 1，其中传统 iGluRs 基因与 *IR8a* 和 *IR25a* 基因的 ω 平均值接近，分别为 0.0623 和 0.1038，说明它们在强的纯化选择下进化，可能编码功能性的受体。而 IRs 基因 ω 的平均值为 0.2110，在较弱的纯化选择下进化，因此可能包含更多的正向选择位点，执行不同的功能。

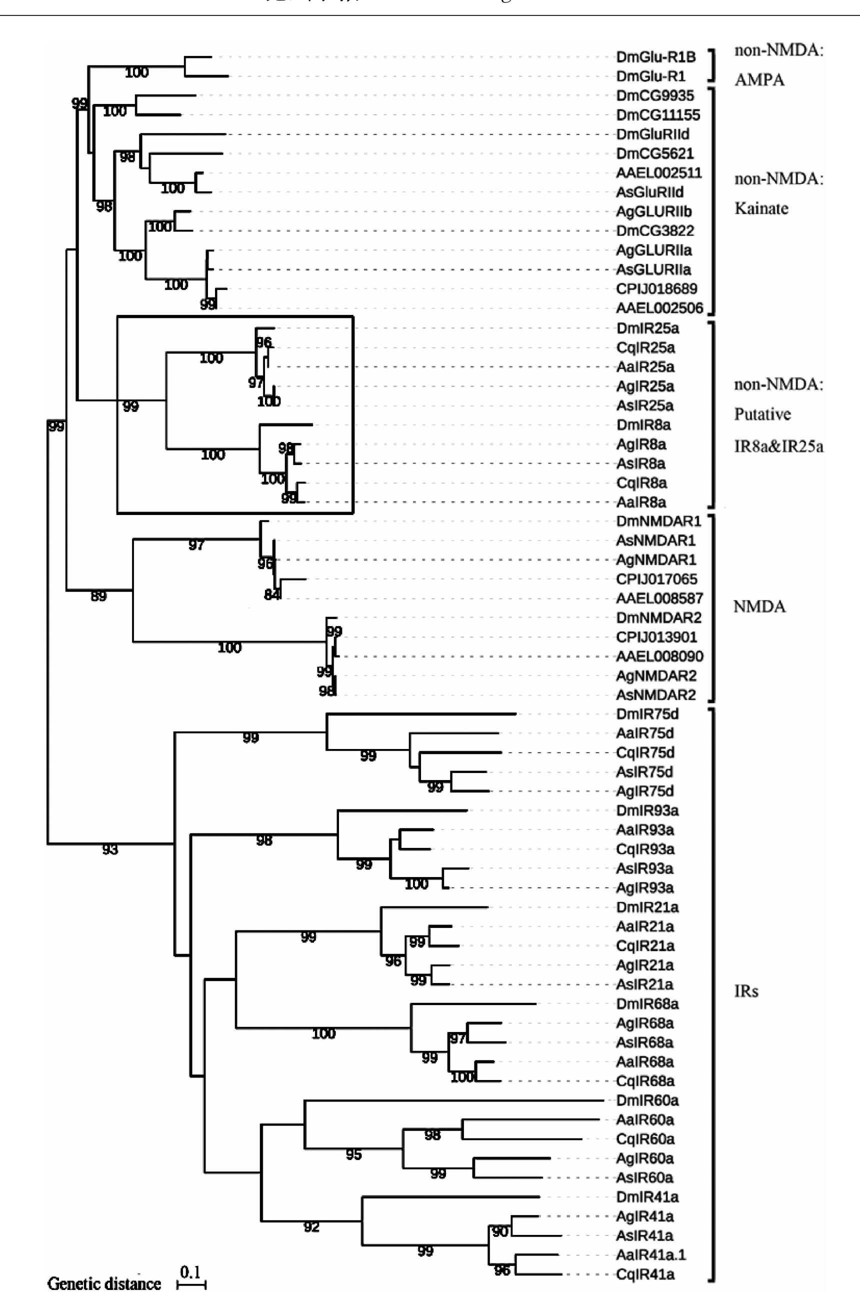


图 5 基于氨基酸序列的 5 种昆虫代表性 iGluRs 和 IRs 的最大似然法(ML)系统发育树

Fig. 5 The phylogenetic relationships of representative iGluRs and IRs based on amino acid sequences in five insect species (maximum likelihood method)

As：中华按蚊 *Anopheles sinensis*；Cq：致倦库蚊 *Culex quinquefasciatus*；Aa：埃及伊蚊 *Aedes aegypti*；Ag：冈比亚按蚊 *Anopheles gambiae*；Dm：黑腹果蝇 *Drosophila melanogaster*. 大于 50% 的 bootstrap 值(1 000 次重复)标记在分支节点上，标尺代表系统遗传距离。不同类型的代表性 iGluRs 和 IRs 基因包括进该系统发育分析中用于显示 3 类基因间的关系。The bootstrap values (1 000 replicates) larger than 50% are marked on the nodes of the phylogenetic tree, and the scale bar indicates the genetic distance. Different types of representative iGluRs and IRs genes are included in the analysis to show the relationships of three types of genes.

3 讨论

离子受体(IRs)是一类高度分化的离子谷氨酸受体(iGluRs)，与嗅觉相关，是化学感应受体中一个新的种类，在检测外界环境及细胞间化学信号的过程中有广泛的功能(Benton, 2007)。在之前的研究中，由于 *IR8a* 与 *IR25a* 基因在触角中表达，可能与嗅觉相关，因此被分为 IR 基因，但是并不能解释为什么它们在结构域、系统发育关系和遗传距离等方

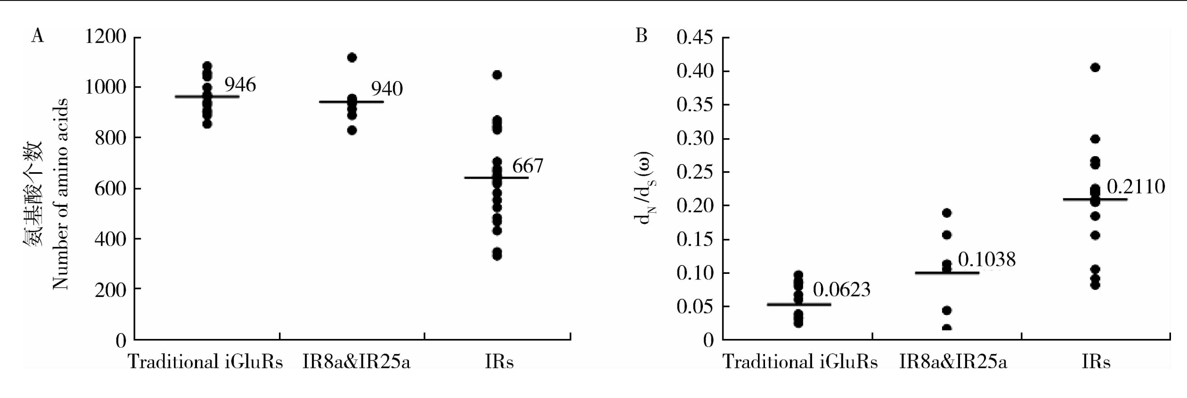


图 6　传统 iGluRs，IR8a 和 IR25a，以及 IRs 氨基酸序列长度变异（A）和 $d_N/d_S(\omega)$ 值（B）比较

Fig. 6　Comparisons of amino acid length variation (A) and $d_N/d_S(\omega)$ values (B) in traditional iGluRs, IR8a & IR25a and IRs

面与传统 iGluRs 更为相近，因此 *IR8a* 与 *IR25a* 的分类地位一直是一个未解决的问题。本研究用生物信息学方法鉴定了 4 种蚊虫 *IR8a* 与 *IR25a* 基因以及代表性的传统 iGluRs 和 IRs 基因。通过比较这 3 类基因在氨基酸序列长度、结构域、配体结合位点保守性和选择压力这 4 个方面的特征可以看出，IR8a 和 IR25a 更接近传统的 iGluRs；从特异性位点和系统发育关系这两个方面的特征可以看出，IR8a 和 IR25a 更接近传统 iGluRs 中的 non-NMDA 受体。因此本研究确定了 IR8a 和 IR25a 的特殊性，将其重新定位为 non-NMDA 类型的受体，并命名为 Putative 受体。

从氨基酸序列长度的平均值看，IR8a 和 IR25a 与传统的 iGluRs 明显更接近，因此可能有更相似的功能。之前的研究表明，iGluRs 的 non-NMDA（Kainate 和 AMPA）和 NMDA 亚类在序列长短上的分化会使其不断适应突触交流机制，从而适应复杂的神经元系统（Croset *et al*., 2010）。其次，从结构域和 3D 结构上看，IR25a 有与传统 iGluRs 一样负责通道的组装以及结合辅助因子的氨基酸末端区域（ATD），IR8a 同样也有类似的区域，只是分散性较强，没有预测到；IRs 则没有 ATD 区，这与之前的研究相符合（Croset *et al*., 2010）。所有传统的 iGluRs 都具有 3 个直接结合谷氨酸的氨基酸位点（R，T 和 D/E），IR8a 同时具有这 3 个位点，IR25a 有精氨酸（R）和天冬氨酸（D），而 IRs 配体结合位点极度分散，3 个位点都没有。此结果与之前的研究结果一致，即黑腹果蝇、豌豆蚜、赤拟谷盗、灰翅夜蛾和家蚕 IRs 配体结合区非常分散，有可能结合不同配体，执行不同功能（Benton *et al*., 2009；郭梅，2014）。同时，我们通过计算非同义突变率与同义突变率的比值（d_N/d_S，ω）来判断蛋白质编码基因的保守程度以及其是否存在选择压力。一般认为，同义突变不受自然选择，而非同义突变受到自然选择的作用，若 $\omega \ll 1$，表示所研究的基因属于纯化选择，$\omega \gg 1$ 表示基因正向选择，当 $\omega = 1$ 时，表示基因在进行中性选择（Swanson and Vacquier, 2002；Swanson *et al*., 2003）。本研究结果发现，*IR8a* 和 *IR25a* 与传统 iGluRs 基因 ω 的平均值更接近，远小于 1，它们在强烈的纯化选择下进化；而 IRs 基因 ω 的平均值较大，在较弱的纯化选择下进化。证明 *IR8a* 和 *IR25a* 基因与传统 iGluRs 基因一样，保守性强，可能编码功能性的受体。另外，IR8a 和 IR25a 与传统 iGluRs 都属于胞质结合蛋白 2 超家族的成员，它们的氨基酸序列与 iGluRs 有相似的功能性特异位点，其中单独与 non-NMDA 受体相似的位点有 12 个，但没有与 NMDA 受体单独相似的位点，因此 IR8a 与 IR25a 更接近传统 iGluRs 中的 non-NMDA 受体。本研究首次利用了离子谷氨酸受体氨基酸特异性位点这个特征来确定基因之间的相似性，并根据位点相似性来进行分类，这为 IR8a 与 IR25a 的重新定位提供了直接证据。系统发育分析也表明，IR8a 和 IR25a 与 non-NMDA 聚在一支上，遗传距离也与其相近，远小于 IRs。这与之前黑腹果蝇（Bengtsson *et al*., 2012；Liu *et al*., 2012）、螟虫和棉铃虫（Liu *et al*., 2012；Glaser *et al*., 2013）等物种离子谷氨酸受体的系统发育关系结果一致。

本研究从受体的结构域、特异性位点和系统发育关系等各个方面比较结果表明，IR8a 和 IR25a 与传统 iGluRs 更为接近，因此可能有与其一样在神经元中发挥作用的功能，这可为以后的相关实验验证工作提供假设。本研究丰富了蚊虫离子谷氨酸受体的基因数据，并有利于推动离子谷氨酸受体基因在蚊虫生长发育中基因的表达、调控、功能及其他生物学方面的研究。

参考文献（References）

Ashby MC, Daw MI, Isaac JTR, Petralia RS, Wenthold R, Contractor A, Swanson GT, Yuzaki M, Pelkey KA, Mcbain CJ, 2008. The Glutamate Receptors. Humana Press, Totowa, NJ.

Bengtsson JM, Trona F, Montagne N, Anfora G, Ignell R, Witzgall P, Jacquinjoly E, 2012. Putative chemosensory receptors of the codling moth, *Cydia pomonella*, identified by antennal transcriptome analysis. *PLoS ONE*, 7(2): e31620.

Benton R, 2007. Sensitivity and specificity in *Drosophila* pheromone perception. *Trends Neurosci*., 30(10): 512–519.

Benton R, Vannice KS, Gomez-Diaz C, Leslie BV, 2009. Variant ionotropic glutamate receptors as chemosensory receptors in *Drosophila*. *Cell*, 136(1): 149–162.

Brockie PJ, Maricq AV, 2003. Ionotropic glutamate receptors in *Caenorhabditis elegans*. *Neurosignals*, 12(3): 108–125.

Chen B, Zhang YJ, He ZB, Li WS, Si FL, Tang Yao, He QY, Qiao L, Yan ZT, Fu WB, 2014. *De novo* transcriptome sequencing and sequence analysis of the malaria vector *Anopheles sinensis* (Diptera: Culicidae). *Parasit. Vectors*, 7(1): 314.

Croset V, Rytz R, Cummins SF, Budd A, Brawand D, Kaessmann H, Gibson TJ, Benton R, 2010. Ancient protostome origin of chemosensory ionotropic glutamate receptors and the evolution of insect taste and olfaction. *PLoS Genet*., 6(8): e1001064.

Glaser N, Gallot A, Legeai F, Montagné N, Poivet E, Harry M, Calatayud PA, Jacquin-Joly E, 2013. Candidate chemosensory genes in the stemborer *Sesamia nonagrioides*. *Int. J. Biol. Sci*., 9(5): 481–495.

Guo M, 2014. Identification and Function Exploration of Ionotropic Receptors and Pheromone Binding Protein in Locust. PhD Dissertation, China Agricultural University, Beijing. [郭梅, 2014. 蝗虫离子型受体和信息素结合蛋白的鉴定及功能研究. 北京: 中国农业大学博士学位论文]

Heidelberg SB, 2004. Ionotropic Glutamate Receptors. Springer-Verlag, Berlin Heidelberg.

Jin RS, Singh SK, Gu S, Furukawa H, Sobolevsky AI, Zhou J, Yan Jin, Gouaux E, 2009. Crystal structure and association behaviour of the GluR2 amino-terminal domain. *EMBO J*., 28(12): 1812–1823.

Kuner T, Seeburg PH, Guy HR, 2003. A common architecture for K^+ channels and ionotropic glutamate receptors? *Trends Neurosci*., 26(1): 27–32.

Liu C, Pitts RJ, Bohbot JD, Jones PL, Wang G, Zwiebel LJ, 2010. Distinct olfactory signaling mechanisms in the malaria vector mosquito *Anopheles gambiae*. *PLoS Biol*., 8(8): 697–704.

Liu NY, Xu W, Papanicolaou A, Dong SL, Anderson A, 2014. Identification and characterization of three chemosensory receptor families in the cotton bollworm *Helicoverpa armigera*. *BMC Genomics*, 15(1): 1–13.

Liu Y, Gu SH, Zhang YJ, Guo YJ, Wang GR, 2012. Candidate olfaction genes identified within the *Helicoverpa armigera* antennal transcriptome. *PLoS ONE*, 7(10): e48260.

Madden DR, 2002. The structure and function of glutamate receptor ion channels. *Nat. Rev. Neurosci*., 3(2): 91–101.

Masse NY, Turner GC, Jefferis GSXE, 2009. Olfactory information processing in *Drosophila*. *Curr. Biol*., 19(16): R700–R713.

Mayer ML, Armstrong N, 2004. Structure and function of glutamate receptor ion channels. *Annu. Rev. Physiol*., 66: 161–181.

Mayer ML, Ghosal A, Dolman NP, Jane DE, 2006. Crystal structures of the kainate receptor GluR5 ligand binding core dimer with novel GluR5-selective antagonists. *J. Neurosci*., 26(11): 2852–2861.

Montell C, 2009. A taste of the *Drosophila* gustatory receptors. *Curr. Opin. Neurobiol*., 19(4): 345–353.

Naur P, Hansen KB, Kristensen AS, Dravid SM, Pickering DS, Olsen L, Vestergaard B, Egebjerg J, Gajhede M, Traynelis SF, 2007. Ionotropic glutamate-like receptor delta2 binds D-serine and glycine. *Proc. Natl. Acad. Sci. USA*, 104(35): 14116–14121.

Swanson WJ, Nielsen R, Yang Q, 2003. Pervasive adaptive evolution in mammalian fertilization proteins. *Mol. Biol. Evol*., 20(1): 18–20.

Swanson WJ, Vacquier VD, 2002. The rapid evolution of reproductive proteins. *Nat. Rev. Genet*., 3(2): 137–144.

Tang Y, Qiao L, Zhang YJ, Che YF, Hong R, Chen B, 2014. Identification and bioinformatics analysis of genes of the CYP6Y subfamily in *Anopheles sinensis* (Diptera: Culicidae). *Acta Entomol. Sin*., 57(6): 663–672. [唐尧, 乔梁, 张玉娟, 车燕飞, 洪瑞, 陈斌, 2014. 中华按蚊 CYP6Y 亚家族基因鉴定和生物信息学分析. 昆虫学报, 57(6): 663–672]

Thompson JD, Gibson TJ, Higgins DG, 2002. Multiple sequence alignment using ClustalW and ClustalX. In: Current Protocols in Bioinformatics, Chapter 2 (Unit 2). 2.3.1–2.3.22.

Tikhonov DB, Magazanik LG, 2009. Origin and molecular evolution of ionotropic glutamate receptors. *Neurosci. Behav. Physiol*., 39(8): 763–773.

Vosshall LB, Stocker RF, 2007. Molecular architecture of smell and taste in *Drosophila*. *Annu. Rev. Neurosci*., 30: 505–533.

Wiens JJ, Servedio MR, 1998. Phylogenetic analysis and intraspecific variation: performance of parsimony, likelihood, and distance methods. *Syst. Biol*., 47(2): 228–253.

Yang Z, 1997. PAML: a program package for phylogenetic analysis by maximum likelihood. *Comput. Appl. Biosci*., 13(13): 555–556.

（责任编辑：马丽萍）